RAW SEQUENCE LISTING
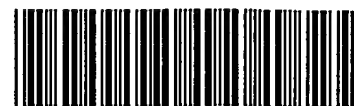
The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: _09/936,367C_
Source: _/Fw/6_
Date Processed by STIC: _3/14/07_

ENTERED

RAW SEQUENCE LISTING   DATE: 03/14/2007
PATENT APPLICATION: US/09/936,367C   TIME: 10:20:45

Input Set : F:\ll\112843-29.ST25.txt
Output Set: N:\CRF4\03142007\I936367C.raw

```
 3 <110> APPLICANT: AFFOLTER, Michael
 4        DE REU, Johannes
 5        VAN DEN BROEK, Peter
 7 <120> TITLE OF INVENTION: EXPRESSION OF PROTEOLYTIC ENZYMES IN KOJI MOLD IN THE
PRESENCE OF
 8        CARBON SOURCES
10 <130> FILE REFERENCE: 112843-29
12 <140> CURRENT APPLICATION NUMBER: US 09/936,367C
13 <141> CURRENT FILING DATE: 2002-01-23
15 <150> PRIOR APPLICATION NUMBER: EP 99 104 923.0
16 <151> PRIOR FILING DATE: 1999-03-11
18 <150> PRIOR APPLICATION NUMBER: PCT/EP00/01796
19 <151> PRIOR FILING DATE: 2000-03-02
21 <160> NUMBER OF SEQ ID NOS: 5
23 <170> SOFTWARE: PatentIn version 3.2
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 4238
27 <212> TYPE: DNA
28 <213> ORGANISM: Aspergillus oryzae
30 <400> SEQUENCE: 1
31 ctgcagttcc agtttctacc ccgtaaatcc ctatcaactt agtccgcccc acattctttt      60
33 ttttttttcc ttttttttc gctcccggtc agagtgatag tgggatttat tacacaccgt     120
35 gcgtggtcga agaacgacac ggaagaagcc ccggaagacg ccttctctag gcaacaaatg     180
37 attgtactct tatgatactc aatacggtag aaaatagaga attgagatac gaaagctgac     240
39 tcatcagaac agaataaggg gaatttttga ttagcaaata acaataataa ttatacaaaa     300
41 aaacaaataa aaaaatttag gggactcccc acccgctgta atcctgggtg tatctcaaag     360
43 caaagcaggc gatctggggg gagcacgttc ttttttttc ttttctcttt tttctatttt     420
45 tttttttttt tttattttag gtctatgcct ttttttttct tttccttttt ttttttttt     480
47 tttgccccc gataattctc cccacacata ggacatactt ttttttttt tccttccact     540
49 cccttcaagg tctccgattc cgataacccc ctctaccagt tcgccctgcc ttttctctc     600
51 ccctcccccg aagctccatt tctctcttct tcccctccat tcctcattct tcctcttccg     660
53 tatttccttt atatgctcct atccccagac catttctcca gatttctctc tctttcccct     720
55 ctctcccttt cgacaaattg ttgcttgact acatccatct cgggttacct acttacagta     780
57 ccaattccgg atatactcta tcccacccat caccacattc cataacagcg ccctttcatt     840
59 gggaaagtca ctcttccttg aaattggtta catcgcggac catcgtacct tctttaatcg     900
61 caaggcttgt gatactcttg cggtgctcgt tcatcaacta gtactttgcc aagagcaagt     960
63 ctccgtcttg tcgggtggtg atcgactctc cccgatttac ctacccctgt gcgacgaat    1020
65 cctgattcgc ctcggctcgt cagcccttcc gagcttccct taagtacagg cttcgtcccc    1080
67 tctttagctg cactcctcgg tgctaggtta ggacgagtca catgccacca ccggcttctt    1140
69 cagtggattt caccaatctg ctgaaccctc agaataacga gactggttct gcaccttcca    1200
71 cgccagtgga tagctccaag gctccctcta ccccgtccag tactcagtcc aactctacca    1260
73 tggcctcgtc tgttagctta ctaccgcccc tcatgaaggg tgctcgtccc gcaacggaag    1320
75 aagcgcgcca ggatcttccc cgtccataca agtgtcccct gtgtgatcgc gccttccatc    1380
77 gtttggagca ccagaccaga catattcgca cacatacggg tgaaaagcca cacgcttgcc    1440
```

```
 79 agttcccggg ctgcacaaaa cgctttagtc gctctgacga gctgacacgc cactcaagaa  1500
 81 ttcacaacaa ccccaactcc aggcggagta acaaggcaca tctggccgct gccgctgccg  1560
 83 ctgccgctgc cggacaagag aatgcaatgg taaatgtgac caacgcgggc tcgttgatgc  1620
 85 ccccgcccac aaagcctatg acccgctctg cgcctgtctc tcaggttgga tctccggatg  1680
 87 tctcccctcc gcactccttc tcgaactatg ccggtcacat gcgttccaat ctgggaccat  1740
 89 atgctcgcaa caccgagcgg gcgtcctcgg gaatggatat caatctactt gccaccgctg  1800
 91 catctcaggt tgagcgtgat gaacaacatt ttgggttcca cgctggtcca cgtaatcacc  1860
 93 atttgttcgc ctcgcgtcac cacaccggtc gtggcctgcc ttccctttca gcgtacgcca  1920
 95 tctcgcacag catgagccgt tctcactttc acgaggacga ggatggttac actcatcgcg  1980
 97 tcaagcgctc aaggcctaac tcaccaaact cgaccgctcc gtcctcaccg actttctctc  2040
 99 acgactctct ttccccaacg ccagaccaca ctccgttggc aacccctgct cattcgccac  2100
101 gcttgaggtc attgggatct agcgaactcc accttccttc gattcgccat ctgtccctcc  2160
103 atcacacccc tgcccttgct ccaatggagc cccagccgga aggccccaac tattacagtc  2220
105 ccagccagtc tcatggtccc acaatcagcg atatcatgtc cagacccgac ggaacacagc  2280
107 gtaaactgcc cgttccacag gttcccaagg tcgcggtgca agatatgctg aaccccagcg  2340
109 ctgggttttc gtcggtttcc tcatcgacga ataactctgt cgcaggaaat gatttggcag  2400
111 aacgtttcta gcctggtgcg gctgcgaaac cctttcaatg tataaagttt tgggctcaaa  2460
113 aaaaattctt gactgtcata cgcgctacga aacgaataga ctttgtgcat ttacagtgcg  2520
115 tggttcatgg gcatccttgg tgtcggctgg ctttcttttgc ttactttgtt cgagtatact  2580
117 tttgcgaggc gtccatagtg atagacgggt gggatattct tgtggctttt tccgtgcttg  2640
119 ttcgattctc ccctttcgct ctccttgaaa aataccttttc ttatcctata accatttgtt  2700
121 tcattatccc aatgggaatt ggctctacag ctcttattca ttttgtctac tcctctcctg  2760
123 aggcccagtc ccctgataat tccgggctct accatataca tttcatttcg actatgtcag  2820
125 tttccgcttc gatttagacc tcgagcagga cgagagggtt ccgaaagaaa atacaaacaa  2880
127 aaattatagt aatctgcgtt tactttggca taatacagta gtcattagtt gaggtaggca  2940
129 taatctggat gtctaaccat cacttgccct aacctcctac catctgctgc tagtatttgt  3000
131 cttacccgaa acccaattca acgagataga tggattgacg aataacaatt tgttgtccag  3060
133 cgacatgcat gatacatgcg tacgtacata cactaatagt agtcacagac cagttcatca  3120
135 catcctggtc tcgggtattc agatacggaa atgcgtaaga ttggaagggt ctaagaaaaa  3180
137 gcaaagaaaa aggaaaagtt aacactggct ggcgctctct ttccatctct gatcaatgtt  3240
139 attgttcgtc actcagctgt ggacgtggct ccagtcaagt tgtgaattat gatagggtat  3300
141 tgttgacttg acaagttgat cttgatggaa tcaaatcttc tccccgccag attctgacgc  3360
143 ttgaggctct cggatcgaat gaacaacttt tcgcaccaca tcaaccggtt gccgcgtgat  3420
145 gctggagaca aaccgaccca aacgtcacgg tcacacggag gatacgtttg ctagagccag  3480
147 ctgatacccc aagagacaag aaggtaaagg tcgcaaaaat cttttcaata agatggcatc  3540
149 ttcccccac caaccttaa ccattctcct ttcaagctgt gttgccccgc tttggtgcat  3600
151 gggcttgggt agtgcggtcg caaaactact aatttaatga ccgactgctg ctgctttttc  3660
153 actcgccgct cacggactaa gcatgtggga acaggatcgc cccgtcacta tttcagatcg  3720
155 tgtcgtatca aggtgttcgc ccggtgctgc tggcacgaac gccggccatc aagatcatt  3780
157 gttctcattc aaaccgggcg gcttacgtct agccgcggac gtaagcacga agagtgtgtg  3840
159 tagtggtggg agtgaagccg ttgccgaaac catgccgtcc tccacggccg tcccgtcgtt  3900
161 atcaagcgac gctgcctccg cttcatcctc atcagcgggt gtatctctgg agacaagatg  3960
163 ggcggaaggt ctcaccggcc aggagatatt agaagacgat ggaacgggcg cgctcgtcgt  4020
165 cccgccgtcc cgccctgctc ggcaatatca tcaccatacc tatatctgtc tgttctatat  4080
167 cttagattgt caccacacct tcgacgatgt cgagcaatgg aagactcacg ttctgagcca  4140
169 cttccgaacc cacgaaccac cgcgaacagc ccgatgccct ctatgtccgg gtgagcggtt  4200
171 cagcgacacc cccgaacaga aaggatggga tcgcatgc                          4238
174 <210> SEQ ID NO: 2
175 <211> LENGTH: 431
```

RAW SEQUENCE LISTING   DATE: 03/14/2007
PATENT APPLICATION: US/09/936,367C   TIME: 10:20:45

Input Set : F:\11\112843-29.ST25.txt
Output Set: N:\CRF4\03142007\I936367C.raw

176 <212> TYPE: PRT
177 <213> ORGANISM: Aspergillus oryzae
179 <400> SEQUENCE: 2
181 Met Pro Pro Pro Ala Ser Ser Val Asp Phe Thr Asn Leu Leu Asn Pro
182 1   5   10   15
185 Gln Asn Asn Glu Thr Gly Ser Ala Pro Ser Thr Pro Val Asp Ser Ser
186   20   25   30
189 Lys Ala Pro Ser Thr Pro Ser Ser Thr Gln Ser Asn Ser Thr Met Ala
190   35   40   45
193 Ser Ser Val Ser Leu Leu Pro Pro Leu Met Lys Gly Ala Arg Pro Ala
194   50   55   60
197 Thr Glu Glu Ala Arg Gln Asp Leu Pro Arg Pro Tyr Lys Cys Pro Leu
198 65   70   75   80
201 Cys Asp Arg Ala Phe His Arg Leu Glu His Gln Thr Arg His Ile Arg
202   85   90   95
205 Thr His Thr Gly Glu Lys Pro His Ala Cys Gln Phe Pro Gly Cys Thr
206   100   105   110
209 Lys Arg Phe Ser Arg Ser Asp Glu Leu Thr Arg His Ser Arg Ile His
210   115   120   125
213 Asn Asn Pro Asn Ser Arg Arg Ser Asn Lys Ala His Leu Ala Ala Ala
214   130   135   140
217 Ala Ala Ala Ala Ala Ala Gly Gln Gly Gln Glu Asn Ala Met Val Asn
218 145   150   155   160
221 Val Thr Asn Ala Gly Ser Leu Met Pro Pro Thr Lys Pro Met Thr
222   165   170   175
225 Arg Ser Ala Pro Val Ser Gln Val Gly Ser Pro Asp Val Ser Pro Pro
226   180   185   190
229 His Ser Phe Ser Asn Tyr Ala Gly His Met Arg Ser Asn Leu Gly Pro
230   195   200   205
233 Tyr Ala Arg Asn Thr Glu Arg Ala Ser Ser Gly Met Asp Ile Asn Leu
234   210   215   220
237 Leu Ala Thr Ala Ala Ser Gln Val Glu Arg Asp Glu Gln His Phe Gly
238 225   230   235   240
241 Phe His Ala Gly Pro Arg Asn His His Leu Phe Ala Ser Arg His His
242   245   250   255
245 Thr Gly Arg Gly Leu Pro Ser Leu Ser Ala Tyr Ala Ile Ser His Ser
246   260   265   270
249 Met Ser Arg Ser His Phe His Glu Asp Glu Asp Gly Tyr Thr His Arg
250   275   280   285
253 Val Lys Arg Ser Arg Pro Asn Ser Pro Asn Ser Thr Ala Pro Ser Ser
254   290   295   300
257 Pro Thr Phe Ser His Asp Ser Leu Ser Pro Thr Pro Asp His Thr Pro
258 305   310   315   320
261 Leu Ala Thr Pro Ala His Ser Pro Arg Leu Arg Ser Leu Gly Ser Ser
262   325   330   335
265 Glu Leu His Leu Pro Ser Ile Arg His Leu Ser Leu His His Thr Pro
266   340   345   350
269 Ala Leu Ala Pro Met Glu Pro Gln Pro Glu Gly Pro Asn Tyr Tyr Ser
270   355   360   365

```
273 Pro Ser Gln Ser His Gly Pro Thr Ile Ser Asp Ile Met Ser Arg Pro
274       370               375             380
277 Asp Gly Thr Gln Arg Lys Leu Pro Val Pro Gln Val Pro Lys Val Ala
278 385               390             395             400
281 Val Gln Asp Met Leu Asn Pro Ser Ala Gly Phe Ser Ser Val Ser Ser
282               405             410             415
285 Ser Thr Asn Asn Ser Val Ala Gly Asn Asp Leu Ala Glu Arg Phe
286           420             425             430
289 <210> SEQ ID NO: 3
290 <211> LENGTH: 29
291 <212> TYPE: DNA
292 <213> ORGANISM: Artificial
294 <220> FEATURE:
295 <223> OTHER INFORMATION: Oligonucleotide
297 <400> SEQUENCE: 3
298 cttccccgtc catagtagtg tcccctgtg                              29
301 <210> SEQ ID NO: 4
302 <211> LENGTH: 29
303 <212> TYPE: DNA
304 <213> ORGANISM: Artificial
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Oligonucleotide
309 <400> SEQUENCE: 4
310 cacaggggac actactatgg acggggaag                              29
313 <210> SEQ ID NO: 5
314 <211> LENGTH: 6
315 <212> TYPE: PRT
316 <213> ORGANISM: Artificial
318 <220> FEATURE:
319 <223> OTHER INFORMATION: consensus of creA DNA-binding site
321 <400> SEQUENCE: 5
323 Ser Tyr Gly Arg Gly Gly
324 1               5
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 03/14/2007
PATENT APPLICATION: **US/09/936,367C**   TIME: 10:20:46

Input Set : **F:\ll\112843-29.ST25.txt**
Output Set: **N:\CRF4\03142007\I936367C.raw**

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:3,4,5

VERIFICATION SUMMARY   DATE: 03/14/2007
PATENT APPLICATION: US/09/936,367C   TIME: 10:20:46

Input Set : F:\11\112843-29.ST25.txt
Output Set: N:\CRF4\03142007\I936367C.raw